QUOL: Qualitative, Objective Likelihoodism

Conor Mayo-Wilson, Soham Pardeshi, and Aditya Saraf

University of Washington

FEW
June 30, 2020

Bernstein on Blood Types

Question: How do genes determine blood type?

H_1 : Single-locus hypothesis.
Three alleles A, B, and O determine type.

H_2 : Two locus hypothesis.
Two alleles A/a and B/b determine type.

Bernstein's Data

H_1 : Single Locus Hypothesis

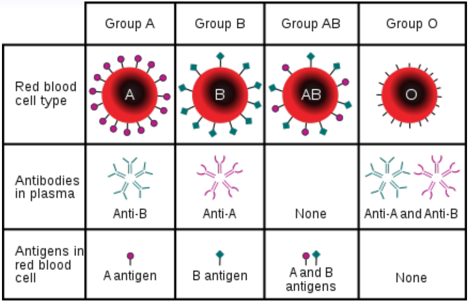
Type	Genotypes	Observed %	Expected % under H_1
A	AA, AO	.422	.4112
B	BB, BO	.206	.1943
AB	AB	.078	.0911
O	OO	.294	.3034

H_2 : Two Locus Hypothesis

Type	Genotypes	Observed %	Expected % under H_2
A	AAbb, Aabb	.422	.358
B	aaBB, aaBb	.206	.142
AB	AABB,AaBB, AABb, AaBb	.078	.142
O	aabb	.294	.358

Sample Size = 502

Bernstein and Law of Likelihood



According to Edwards, Bernstein's data is about 400 million times more likely if H_1 is true than if H_2 is!

$$P_{H_1}(E) \approx 4 \cdot 10^8 \cdot P_{H_2}(E)$$

where E is Bernstein's data.

Law of Likelihood

[Edwards, 1984] uses this example to motivate . . .

Law of Likelihood

Law of Likelihood (LL): E favors θ_1 to θ_2 if and only if

$$P_{\theta_1}(E) > P_{\theta_2}(E)$$

Notes:

- Above, θ_1 and θ_2 are *simple* (i.e., not composite) statistical hypotheses.
- As we use the term, LL is strictly weaker than the thesis that the likelihood ratio is a numerical strength of evidence.

Likelihoodism: What is it?

Likelihoodism is often glossed as the thesis that "all the information that about an unknown parameter is contained in the likelihood function."

In practice, it's taken to be the conjunction of LL and the *likelihood principle* (LP), which is best illustrated by an uncontroversial consequence . . .

Conditionality Principle

- Suppose Gertrude conducts experiment $\mathbb{E}$ and gets data E.
- Allan can't decide whether to conduct experiment $\mathbb{E}$ or $\mathbb{F}$. He flips a coin and decides to conduct experiment $\mathbb{E}$ too. He gets the exact same data E as Gertrude.
- **Conditionality Principle:** Gertrude and Allan have equivalent evidence.

Conditionality Principle and LP

The conditionality principle follows from the likelihood principle . . .

Likelihood Principle

Experiment 1

	Observe A	Observe B
θ_1	10%	90%
θ_2	20%	80%
θ_3	5%	95%

Experiment 2

	Observe D	...
θ_1	20%	...
θ_2	40%	...
θ_3	10%	...

Likelihood Principle

Experiment 1

	Observe A	Observe B
θ_1	10%	90%
θ_2	20%	80%
θ_3	5%	95%

Experiment 2

	Observe D	...
θ_1	20%	...
θ_2	40%	...
θ_3	10%	...

Likelihood Principle (LP): For any two statistical experiments $\mathbb{E}$ and $\mathbb{F}$ and any pair of observations E and F from $\mathbb{E}$ and $\mathbb{F}$ respectively, if there is some constant $c > 0$ such that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all simple hypotheses $\theta \in \Theta$, then E and F are evidentially equivalent.

LP entails Conditionality

Example:

- Let $\mathbb{F}$ be the experiment in which one flips a coin of *known* bias $c > 0$ to decide whether to conduct $\mathbb{E}$ or $\mathbb{E}'$.
- Then the outcome $F = \langle \text{Heads}, E \rangle$ of $\mathbb{F}$ is always c times the probability of obtaining E if one had simply conducted the experiment $\mathbb{E}$.
- So E and $F = \langle \text{Heads}, E \rangle$ are evidentially equivalent by $\textsc{lp}$.

We have two overarching goals . . .

Favoring, Equivalence, and Objectivity

Goal 1: Clarify the importance of "favoring" and "evidential equivalence"

- If "favoring" is not pre-theoretic and does not tell us what to believe or how to act, why explicate it?
- In what sense are LL and LP **objective**?

Likelihoodism: A complete theory

Goal 2: Provide a **complete** likelihoodist theory of evidence

Likelihoodism: A complete theory

Goal 2: Provide a **complete** likelihoodist theory of evidence

1. Extend LL to composite hypotheses, including those with nuisance parameters.

Likelihoodism: A complete theory

Goal 2: Provide a **complete** likelihoodist theory of evidence

1. Extend LL to composite hypotheses, including those with nuisance parameters.
2. Characterize **comparative** favoring (i.e., when E favors H_1 over H_2 at least as much as F does) so that the analysis
   1. Unifies LL and LP,
   2. Entails truisms about evidential strength that are not entailed by LL and LP (e.g., larger iid samples are typically stronger evidence)

Likelihoodism: A complete theory

Goal 2: Provide a **complete** likelihoodist theory of evidence

1. Extend LL to composite hypotheses, including those with nuisance parameters.
2. Characterize **comparative** favoring (i.e., when E favors H_1 over H_2 at least as much as F does) so that the analysis
    1. Unifies LL and LP,
    2. Entails truisms about evidential strength that are not entailed by LL and LP (e.g., larger iid samples are typically stronger evidence)
3. Motivate **qualitative** likelihoodist principles that generalize LL, LP, and the comparative principle that unifies them.

Objectivity and Agreement

Likelihoodism and Objectivity

Question: Putting aside the question of what "favoring" and "evidentially equivalent" mean, why do authors often claim that LL and LP provide **objective** standards of evidence?

Answer: LL and LP employ only likelihoods – probabilities of the form $P_\theta(E)$ – to characterize "favoring" and "equivalence." Okay, but why does that matter?

Likelihoodism and Objectivity

Question: Why does employing only likelihoods – probabilities of the form $P_\theta(E)$ – make a principle more "objective"?

The Standard Answer:

Claim 1: P_θ describe **frequencies** or **chances**, i.e., "objective" features of the world that are independent of the experimenter.

- E.g., In the blood-types example, P_θ describes frequencies of alleles in a population.

Likelihoodism and Objectivity

Question: Why does employing only likelihoods – probabilities of the form $P_\theta(E)$ – make a principle more "objective"?

The Standard Answer:

Claim 1: P_θ describe **frequencies** or **chances**, i.e., "objective" features of the world that are independent of the experimenter.

- E.g., In the blood-types example, P_θ describes frequencies of alleles in a population.

Claim 2: Posterior probabilities of the form $P(\theta|E)$ describe a particular agent's credences, which are considered "subjective."

Clarifying Objectivity

Problem: We think the standard answer conflates two notions of objectivity.

1. Objective measures of evidence do not depend upon a **single** agent's credences (i.e., a particular credence function).
2. Objective measures of evidence depend only upon frequencies or chances.

Clarifying Objectivity

Problem: We think the standard answer conflates two notions of objectivity.

1. Objective measures of evidence do not depend upon a **single** agent's credences (i.e., a particular credence function).
2. Objective measures of evidence depend only upon frequencies or chances.

The first claim is reasonable; we reject the second.

Likelihoodism and Objectivity

Our Thesis: LL and LP can be understood as answering questions about how **all** rational agents updates their credences.

In [Douglas, 2009]'s terminology, LL and LP characterize **concordantly objective** updating procedures.

The Key Concept: Comparative Bayesian Favoring

Posteriors and Likelihoods

Definition: Given a prior Q and likelihood functions $\{P_\theta(E)\}_{\theta \in \Theta}$, the **posterior** of $Q(\cdot|E)$ is defined by

$$\begin{aligned} Q(H|E) &= \frac{Q(E|H) \cdot Q(H)}{Q(E)} \\ &= \frac{\sum_{\theta \in H} P_\theta(E) \cdot Q(\theta)}{\sum_{\theta \in \Theta} P_\theta(E) \cdot Q(\theta)} \end{aligned}$$

In other words, we assume agents **agree** to update using common likelihood functions.

Question: Where else will they agree?

Definition: Say that E comparatively Bayesian favors H_1 to H_2 more than F does if

$$Q(H_1|E \cap (H_1 \cup H_2)) \geq Q(H_1|F \cap (H_1 \cup H_2))$$

for all priors such that that $Q(H_1), Q(H_2) > 0$.

Notation: Write $E \, {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$ in this case.

Idea: Learning E raises one's confidence in H_1 at least as much as learning F *no matter one's initial degrees of belief.*

Theorem

Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then E comparatively Bayesian favors H_1 to H_2 at least as much as F iff

$$\frac{\min_{\theta \in H_1} P_\theta^{\mathbb{E}}(E)}{\max_{\theta \in H_2} P_\theta^{\mathbb{E}}(E)} \geq \frac{\max_{\theta \in H_1} P_\theta^{\mathbb{F}}(F)}{\min_{\theta \in H_2} P_\theta^{\mathbb{F}}(F)}.$$

Theorem

Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then E comparatively Bayesian favors H_1 to H_2 at least as much as F iff

$$\frac{\min_{\theta \in H_1} P_\theta^{\mathbb{E}}(E)}{\max_{\theta \in H_2} P_\theta^{\mathbb{E}}(E)} \geq \frac{\max_{\theta \in H_1} P_\theta^{\mathbb{F}}(F)}{\min_{\theta \in H_2} P_\theta^{\mathbb{F}}(F)}.$$

Corollary

Suppose $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple hypotheses. Then E comparatively Bayesian favors H_1 to H_2 at least as much as F iff

$$\frac{P_{\theta_1}^{\mathbb{E}}(E)}{P_{\theta_2}^{\mathbb{E}}(E)} \geq \frac{P_{\theta_2}^{\mathbb{F}}(F)}{P_{\theta_1}^{\mathbb{F}}(F)}.$$

Application: Favoring Larger Samples

- Let E be 75 heads in 100 tosses.
- Let F be three heads in four tosses.
- θ_1 = The coin's bias is $3/4$.
- θ_2 = The coin's bias is a specific value other than $3/4$.

Then E comparatively favors θ_1 to θ_2 more than F does, but not vice versa.

LL and Agreement

Favoring

Definition: Say that E strictly Bayesian favors H_1 to H_2 if

$$Q(H_1|E \cap (H_1 \cup H_2)) > Q(H_1|(H_1 \cup H_2))$$

for all priors such that that $Q(H_1), Q(H_2) > 0$.

Equivalently: E strictly favors H_1 to H_2 more than $F = \Omega$, the sure event.

Theorem

Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then E comparatively Bayesian favors H_1 to H_2 at least as much as F iff

$$\frac{\min_{\theta \in H_1} P_\theta^{\mathbb{E}}(E)}{\max_{\theta \in H_2} P_\theta^{\mathbb{E}}(E)} \geq \frac{\max_{\theta \in H_1} P_\theta^{\mathbb{F}}(F)}{\min_{\theta \in H_2} P_\theta^{\mathbb{F}}(F)}.$$

Theorem

Suppose H_1 and H_2 are finite and disjoint. Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then E comparatively Bayesian favors H_1 to H_2 at least as much as F iff

$$\frac{\min_{\theta \in H_1} P_\theta^{\mathbb{E}}(E)}{\max_{\theta \in H_2} P_\theta^{\mathbb{E}}(E)} \geq \frac{\max_{\theta \in H_1} P_\theta^{\mathbb{F}}(F)}{\min_{\theta \in H_2} P_\theta^{\mathbb{F}}(F)}.$$

Corollary

Suppose H_1 and H_2 are finite. Then E strictly Bayesian favors H_1 to H_2 if and only if

$$\min_{\theta \in H_1} P_\theta^{\mathbb{E}}(E) > \max_{\theta \in H_2} P_\theta^{\mathbb{E}}(E).$$

If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, then E strictly Bayesian favors H_1 to H_2 if and only if $P_{\theta_1}^{\mathbb{E}}(E) > P_{\theta_2}^{\mathbb{E}}(E)$.

Moral: The law of likelihood answers the question: "When will **all** Bayesians' credences in H_1 increase upon learning E, if they assume either H_1 or H_2 is true?"

LP and Agreement

Likelihood Principle

Experiment 1

	Observe A	Observe B
θ_1	10%	90%
θ_2	20%	80%
θ_3	5%	95%

Experiment 2

	Observe D	...
θ_1	20%	...
θ_2	40%	...
θ_3	10%	...

Likelihood Principle (LP): For any two statistical experiments $\mathbb{E}$ and $\mathbb{F}$ and any pair of observations E and F from $\mathbb{E}$ and $\mathbb{F}$ respectively, if there is some constant $c > 0$ such that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all simple hypotheses $\theta \in \Theta$, then E and F are evidentially equivalent.

Bayesian Favoring Equivalence

Definition: Say E and F are **Bayesian-favoring equivalent** if $E \; {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} F$ and $F \; {}_{H_1}^{\mathcal{B}}\!\trianglerighteq_{H_2} E$ for all disjoint hypotheses H_1 and H_2, i.e.,

$$Q(H_1|E \cap (H_1 \cup H_2)) = Q(H_1|F \cap (H_1 \cup H_2))$$

for all disjoint hypotheses H_1 and H_2 and *for all priors* Q such that $Q(H_1), Q(H_2) > 0$.

Theorem

Let E and F be outcomes of experiments $\mathbb{E}$ and $\mathbb{F}$ respectively. Then the following are equivalent:

1. *E and F are Bayesian favoring equivalent,*
2. *$Q(H|E) = Q(H|F)$ for all hypotheses H and priors Q, and*
3. LP *entails E and F are evidentially equivalent, i.e., there is some $c > 0$ such that $P_\theta^{\mathbb{E}}(E) = c \cdot P_\theta^{\mathbb{F}}(F)$ for all θ.*

LP and Bayesian Evidential Equivalence

Morals:

1. LP characterizes **precisely** when every member of the hypothetical community of Bayesian scientists would agree two pieces of evidence are equivalent.
2. The notions of "favoring" and "evidential equivalence" axiomatized by LL and LP respectively can be interpreted as special cases of the notion of comparative Bayesian-favoring.

Further Applications

Fans of Bayesianism:

1. The methodology suggests a way for finding other statistical principles about favoring and equivalence.
   - E.g., When H_1 and H_2 are composite, the question of how to extend LL is debated.
   - E.g., When two experiments $\mathbb{E}$ and $\mathbb{F}$ have different parameter spaces, LP says nothing.

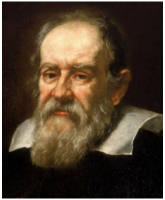
Further Applications

Fans of Bayesianism:

1. The methodology suggests a way for finding other statistical principles about favoring and equivalence.
   - E.g., When H_1 and H_2 are composite, the question of how to extend LL is debated.
   - E.g., When two experiments $\mathbb{E}$ and $\mathbb{F}$ have different parameter spaces, LP says nothing.
2. The "social" perspective (i.e., viewing statistical principles as implicit agreements) tells us when we've found *all* the relevant statistical principles for the Bayesian community.

There are at least two related ways in which we think the Bayesian motivation for likelihoodist principles is limited . . .

Qualitative Evidence



	Follow Jupiter in Retrograde	Don't follow
Orbit Jupiter	Likely	Impossible
Don't orbit	Very unlikely	Extremely likely

In most cases in the history of science, we can't assign precise probabilities to various observations. Yet we seem capable of drawing conclusions about what the evidence favors.

Bayesianism Entails Logical Omniscience

Bayesianism:

1. Bayesianism requires agents to have degrees of belief that obey the probability axioms.
2. Probability axioms: maximal probability and additivity.
3. Probability domain: an algebra of sets (events) or a logically closed set of formulae

Bayesianism Entails Logical Omniscience

Bayesianism:

1. Bayesianism requires agents to have degrees of belief that obey the probability axioms.
2. Probability axioms: maximal probability and additivity.
3. Probability domain: an algebra of sets (events) or a logically closed set of formulae

Logical Omniscience:

1. Agents must be maximally sure of any tautologies.
2. Agents must recognize any logically equivalent formulae/different descriptions of the same set.
   - For sets, this follows from the fact that two sets are equivalent if they have the same elements (and that the probability function is well-defined).
   - For formulae, it can easily be derived from the axioms that logically equivalent formulae should be assigned the same probability.

Why Keep Bayesianism?

The previous argument seems like an equally good argument for not endorsing a Bayesian view of LL/LP. Indeed, if you think nothing more needs to be clarified regarding the notions of favoring and evidential equivalence, there's no need to move to qualitative probability. However, both the unifying and explanatory nature of our Bayesian framework makes it seem more attractive to address these issues without abandoning Bayesian ideas entirely.

QUOL: Qualitative, Objective Likelihoodism

Statistical problems

Typical statistical problems have three components:

- A set Θ of simple hypotheses,
- A set Ω consisting of possible data sequences, and
- For each hypothesis θ, there is a probability measure P_θ over sets of data sequences in Ω.

Statistical problems

Typical statistical problems have three components:

- A set Θ of simple hypotheses,
- A set Ω consisting of possible data sequences, and
- For each hypothesis θ, there is a probability measure P_θ over sets of data sequences in Ω.

Bayesians claim that agents' degrees of belief can be represented by a function Q that assigns numerical probabilities to $\Theta \times \Omega$.

- Often, Bayesians assume $Q(\cdot|H) = P_\theta$ so that the likelihoods are intersubjectively shared.

Our Idea: We relax numerical probability functions to qualitative orderings. This affects

- The likelihood functions, and
- Agents' degrees of belief

Qualitative Likelihoods

Instead of $\{P_\theta\}_{\theta \in \Theta}$, we will model likelihoods using a relation $\sqsubseteq$.

- Informally: $A|\theta \sqsubseteq B|\theta'$ means "B is more likely if θ' is true than A would be if θ were true."

Qualitative Likelihoods

Instead of $\{P_\theta\}_{\theta \in \Theta}$, we will model likelihoods using a relation $\sqsubseteq$.

- Informally: $A|\theta \sqsubseteq B|\theta'$ means "B is more likely if θ' is true than A would be if θ were true."

Notation: $A|\theta \equiv B|\theta'$ is defined as holding when $A|\theta \sqsubseteq B|\theta'$ and vice versa.

Qualitative Priors

Instead of a prior (or posterior) Q, we we will model degrees of belief using a relation $\preceq$.

Qualitative Priors

Instead of a prior (or posterior) Q, we we will model degrees of belief using a relation $\preceq$.

Notation: $H|E \sim H'|F$ is defined as holding when $H|E \preceq H'|F$ and vice versa.

Qualitative Priors

- Let $\mathcal{F}$ be an algebra containing events in $\Theta \times \Omega$ and let $\mathcal{N} \subseteq \mathcal{F}$.
- An agent's beliefs are represented by a binary relation $\preceq$ on $\mathcal{F} \times (\mathcal{F} \setminus \mathcal{N})$ such that
  - C0. $A|H \preceq B|H'$ if and only if $A|H \sqsubseteq B|H'$, i.e., agents agree upon the likelihood of various data *given* the various hypotheses,
  - C1. $\preceq$ is a weak ordering,
  - C2. $A \in \mathcal{N}$ if and only if $A|\Theta \sim \emptyset|\Theta$,
  - C3. $A|A \sim B|B$ for all A, B,
  - C4. $A \cap B|B \sim A|B$,
  - C5. If $A_1|C_1 \preceq A_2|C_2$ and $B_1|C_1 \preceq B_2|C_2$ and $A_1 \cap B_1 = A_2 \cap B_2 = \emptyset$, then $A_1 \cup B_1|C_1 \preceq A_2 \cup B_2|C_2$.
  - C6. Suppose $C \subseteq B \subseteq A$ and $C' \subseteq B' \subseteq A'$. If $B|A \preceq C'|B'$ and $C|B \preceq B'|A'$, then $C|A \preceq C'|A'$.

Axioms C1-C6 are some of [Krantz et al., 2006]'s axioms for qualitative conditional probability.

How Does Qualitative Probability Avoid Logical Omniscience?

Agents are still required to have orderings that agree with some subset of the axioms. However:

1. There is no maximal probability associated with the orderings*. So agents need not be equally or maximally sure of any tautology.
2. When we move from sets/events to formulae, there will be no requirement that logically equivalent formulae have the same probability.

Further Motivation

Frequentists and likelihoodists often don't want to assign probabilities to hypotheses, but it seems like they will admit to having or wanting an ordering of hypotheses.

- Saying that the evidence "favors" one hypothesis or that we ought to reject the null hypothesis seems to imply some sort of ordering.

Note:

- These axioms are not sufficient for $\preceq$ to be representable as a conditional probability function.

The Key Concept: Comparative Qualitative Favoring

Qualitative Comparative Favoring

Definition: Say E comparatively favors H_1 to H_2 at least as much as F if

$$H_1|E \cap (H_1 \cup H_2) \succeq H_1|F \cap (H_1 \cup H_2)$$

for all the orderings such that $H_1|\Theta, H_2|\Theta \succ \emptyset|\Theta$.

Comparative Favoring, Equivalence, and Non-Comparative

As in the quantitative case, we can use this comparative notion to define notions of

- E strictly qualitatively favors H_1 to H_2, and
- E and F are qualitative favoring equivalent

See the handout for precise definitions.

Qualitative LL

Community Favoring

Qualitative Law of Likelihood (QLL): E favors θ_1 to θ_2 if and only if $E|\theta_2 \sqsubset E|\theta_1$.

Theorem

Suppose H_1 and H_2 are finite and disjoint. Then E strictly favors H_1 to H_2 if and only if $E|\theta_2 \sqsubset E|\theta_1$ for all $\theta_1 \in H_1$ and all $\theta_2 \in H_2$.

If $H_1 = \{\theta_1\}$ and $H_2 = \{\theta_2\}$ are simple, then E strictly qualitatively favors H_1 to H_2 if and only if $E|\theta_2 \sqsubset E|\theta_1$, i.e., iff QLL *entails so.*

Qualitative LP

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

By LP, pulling an A is equivalent to pulling a B on the first draw, as $P_\theta(A) = 1/2 \cdot P_\theta(B)$ for any simple hypothesis θ (about the urn type).

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

- Let C_1 and C_2 respectively denote the events that, after replacing the first draw, one draws C_1/C_2 on the second draw.

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

- Let C_1 and C_2 respectively denote the events that, after replacing the first draw, one draws C_1/C_2 on the second draw.
- Then $P_{\theta_1}(B \cap C_1) = P_{\theta_1}(B) \cdot P_{\theta_1}(C_1) = 1/2 \cdot P_{\theta_1}(B) = P_{\theta_1}(A)$ by independence of the draws, if the urn is Type 1.

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

- Let C_1 and C_2 respectively denote the events that, after replacing the first draw, one draws C_1/C_2 on the second draw.
- Then $P_{\theta_1}(B \cap C_1) = P_{\theta_1}(B) \cdot P_{\theta_1}(C_1) = 1/2 \cdot P_{\theta_1}(B) = P_{\theta_1}(A)$ by independence of the draws, if the urn is Type 1.
- Similarly, $P_{\theta_2}(B \cap C_2) = P_{\theta_2}(B) \cdot P_{\theta_2}(C_2) = 1/2 \cdot P_{\theta_2}(B) = P_{\theta_2}(A)$.

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

- Let C_1 and C_2 respectively denote the events that, after replacing the first draw, one draws C_1/C_2 on the second draw.
- Then $P_{\theta_1}(B \cap C_1) = P_{\theta_1}(B) \cdot P_{\theta_1}(C_1) = 1/2 \cdot P_{\theta_1}(B) = P_{\theta_1}(A)$ by independence of the draws, if the urn is Type 1.
- Similarly, $P_{\theta_2}(B \cap C_2) = P_{\theta_2}(B) \cdot P_{\theta_2}(C_2) = 1/2 \cdot P_{\theta_2}(B) = P_{\theta_2}(A)$.
- By LP, because $P_{\theta_1}(C_1) = P_{\theta_2}(C_2) > 0$ and $P_{\theta_i}(A) = P_{\theta_i}(B \cap C_i) = P_{\theta_i}(C_i) \cdot P_{\theta_i}(B)$ for all i, we know A and B are evidentially equivalent.

Motivation

	Number of balls				
	A	B	C_1	C_2	Total
Urn Type 1	15	30	50	5	100
Urn Type 2	10	20	20	50	100

Fact: A and B are equivalent according to LP if there is a set of events $\{C_\theta\}_{\theta \in \Theta}$ such that

1. $P_\theta(A) = P_\theta(B \cap C_\theta)$ for all θ,
2. $P_\theta(C_\theta) = P_{\theta'}(C_{\theta'}) > 0$ for all $\theta, \theta' \in \Theta$, and
3. B and C_θ are independent given P_θ

Qualitative LP

Proposition

E and F are favoring equivalent if there is a set of events $\{C_\theta\}_{\theta\in\Theta}$ such that

1. $E|\theta \equiv F \cap C_\theta|\theta$ *for all* θ*,*
2. $C_\theta|\theta \equiv C_\upsilon|\upsilon$ *for all* θ *and* υ*, and*
3. C_θ *and* F *are qualitatively conditionally independent given* θ *for all* θ

A partial converse is stated on the handout.

Philosophical Morals

Upshots 1

Morals: The type of agreement that LL and LP characterize extends to qualitative settings, i.e., to characterize agreement among agents whose beliefs satisfy relatively few coherence constraints.

- Our work, therefore, provides a better rational reconstruction of qualitative assessments of evidence in science, especially those which predate probability theory,
- And classical statisticians might agree with such coherence constraints ...

Upshots 2

Upshots for Philosophy of Statistics: Qualitative conditions like that provided by the previous theorem **entail** statistical principles (e.g., stopping rule principle) the classical statisticians often reject.

References I

Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.

Edwards, A. W. F. (1984). *Likelihood*. CUP Archive.

Fine, T. L. (2014). *Theories of probability: An examination of foundations*. Academic Press.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (2006). *Foundations of Measurement Volume II: Geometrical, Threshold, and Probabilistic Representations*. Dover Publications, Mineola, N.Y.

Thanks.

Suggestions? Questions? Concerns?

Qualitative Conditional Independence

Let's return to the qualitative setting ...

Definition: Given an ordering $\preceq$ representing an agent's degrees of belief, define: $A \perp_C^{\preceq} B$ to hold if and only if

- $A|B \cap C \sim A|C$, and
- $B|A \cap C \sim B|C$.

The idea is that $A \perp_C^{\preceq} B$ represents the claim that A and B are **conditionally independent** given C.

Qualitative Independence

we must be cognizant of the fact that invocations of [stochastic independence] are usually not founded upon empirical or objective knowledge of probabilities. Quite the contrary. Independence is adduced to permit us to simplify and reduce the family of possible probabilistic descriptions for a given experiment.

[Fine, 2014, p. 80]